intervals HPV-based screening was more effective than cytology alone, with a relative reduction in cervical cancer incidence of 49%-90% compared to 33%-80% with cytology alone (depending on screening intervals). In HPV- compared to cytology screening the incremental gain in effectiveness was higher with extended screening intervals and the increase in harms lower. Based on the BHF, 12 of 17 screening strategies were dominated, including annual cytology, the current recommended standard in Germany. Biennial HPV-screening was similarly effective as annual cytology and reduced unnecessary treatment. Moving from biennial HPV- with cytological triage to annual HPV-screening alone results in an incremental harm-benefit ratio of 15-533 unnecessary treatments per additional prevented cervical cancer case (depending on screening adherence rate). **CONCLUSIONS:** The benefit-harm frontier is a useful tool to demonstrate the trade-off between expected gains and risks of different screening strategies. Based on our analyses, HPV-based cervical cancer screening is more effective than cytology alone, but has a higher risk of overtreatment when used in annual screening. In the German health care context, depending on screening adherence rates biennial or triennial HPV-screening for women ≥30 years is similarly effective as annual cytology with significantly reduced unnecessary treatments.

## PRM2
## EVALUATING WHETHER INCONSISTENCIES ARE PRESENT IN A MIXED TREATMENT COMPARISON OF TROUGH FORCED EXPIRATORY VOLUME IN 1 SECOND AT 12 WEEKS

Buckley F[1], Baldwin M[2], Keininger D[3], Jansen J[1], Cope S[4]
[1]*Mapi, Boston, MA, USA,* [2]*Novartis Pharmaceuticals UK Limited, Horsham, West Sussex, UK,* [3]*Novartis Pharma AG, Basel, Switzerland,* [4]*Mapi, Toronto, ON, Canada*

**OBJECTIVES:** To evaluate whether there are inconsistencies in the network of randomized controlled trials (RCTs) used for a network meta-analysis (NMA) comparing alternative long-acting bronchodilators among patients with moderate to severe chronic obstructive pulmonary disease (COPD) in terms of trough forced expiratory volume in 1 second (FEV1) at 12 weeks. **METHODS:** The change from baseline (CFB) in FEV as observed with placebo, tiotropium 18µg/5µg once daily (OD), salmeterol 50µg twice daily (BID), formoterol 12µg BID, aclidinium 400µg BID, glycopyrronium 50µg OD, indacaterol 75/150/300µg OD, formoterol 12µg+ tiotroipum 18µg BID/OD, indacaterol 150µg+ tiotropium 18µg OD, and indacaterol 110µg+ glycopyrronium 50µg OD in RCTs identified with a systematic literature review were synthesized with a NMA. Where possible, treatment estimates from fixed effect (FE) and random effects (RE) NMA models (assuming consistency between direct and indirect evidence) and independent means (IM) models (pooled direct evidence) were compared to assess whether any inconsistencies in the network were present. **RESULTS:** Thirty-two RCTs identified through a systematic literature review were included in the analysis. Direct evidence was available for the monotherapies versus placebo, the combination therapies versus tiotropium, for indacaterol+ glycopyrronium versus placebo, and for tiotropium versus salmeterol. The largest differences between the estimated treatment effect estimates from the NMA and the IM models were observed for the comparisons between indacaterol 150µg versus tiotropium (FE difference=0.025 [95% Credible Intervals (95%CrI): 0.002, 0.047]; RE difference=0.027 [95%CrI: -0.007, 0.61]), indacaterol+ glycopyrronium versus placebo (FE difference=-0.022 [95%CrI: -0.053, 0.008]; RE difference=-0.018 [95%CrI: -0.059, 0.022]), and indacaterol+ glycopyrronium versus tiotropium (FE difference=0.011 [95%CrI: -0.014, 0.036]; RE difference=0.015 [95%CrI: -0.024, 0.053]). **CONCLUSIONS:** Based on a comparison of the findings of a NMA and IM models, some minor inconsistencies in treatment effects for trough FEV1 at 12 weeks were identified that will be explored through additional sensitivity analyses.

## PRM3
## TESTING THE EUNETHTA INTERNAL VALIDITY OF RANDOMIZED CONTROLLED TRIALS GUIDELINE AND TOOL IN HUNGARY

Huszti Z, Juhasz J
*National Institute for Quality- and Organizational Development in Healthcare and Medicines, Budapest, Hungary*

**OBJECTIVES:** The reliability of the results of a randomized trial depends on the extent to which potential sources of bias have been avoided. We tested the EUnetHTA Internal Validity guideline so that to harmonize our risk of bias assessments with the European standard and finally to improve the reliability of relative efficacy and cost effectiveness assessments for decision makers in Hungary. **METHODS:** We translated the risk of bias standardized assessment questions of the EUnetHTA Internal Validity (of randomized controlled trials) into Hungarian. We first chose ten studies for internal validity testing from the ones that were submitted for reimbursement at the beginning of 2012, and their results were used for health economy assessment. **RESULTS:** We found adequate randomization sequence generation in seven studies and we marked it unclear in three trials (e.g.: lack of information, age related sequence generation). The allocation concealment was labeled suitable in six studies (e.g.: IVRS, IWRS) and unclear in four trials. All studies could be classified according the the type of blinding. We found selective reporting in one trial where the non-inferiority results in the per-protocol population were not published. We rated the risk of bias low for eight trials and high for two trials due to unclear sequence generation and publication bias. We also evaluated 77 endpoints and we labelled 22 endpoints with high risk of bias. The most common reasons for high risk ratings were the not appropriately implemented ITT principle and selective reporting. **CONCLUSIONS:** The EUnetHTA guideline gives an opportunity to estimate the risk of bias of randomized controlled trials in a structured and harmonized way without leaving out any important considerations. The results of the internal validity evaluation can lead the focus of interest to those endpoints where the sensitivity analysis is requisite in the health economic models.

## PRM4
## CLINICAL OUTCOMES ASSESSMENTS IN SCHIZOPHRENIA: A SYSTEMATIC LITERATURE REVIEW

Cooke C[1], Heatley R[2], Galani Berardo C[3], Johnson KI[2], Kasper S[4]
[1]*Complete Clarity, Macclesfield, UK,* [2]*Complete Clarity, Macclesfield, Cheshire, UK,* [3]*F. Hoffman La Roche, Basel, Switzerland,* [4]*Medical University of Vienna, Vienna, Austria*

**OBJECTIVES:** There is a growing interest from health technology assessment agencies in determining the clinical outcomes assessments and endpoint strategies that can establish treatment benefits. We describe a systematic literature review of endpoints and outcomes used in schizophrenia trials to determine treatment benefit. **METHODS:** The therapies selected in the search strategy included pharmacological interventions, cognitive-behavioural therapies, family intervention, and music therapy. These were chosen to reflect the range of interventions in current use, and to allow comparison between outcomes reported for different therapies. The search terms were designed to include all outcomes for each therapy area, and were used to search four electronic databases for published English language studies. Randomised controlled trials (RCTs) were retrieved if they included patients with schizophrenia treated with the chosen therapies, and clinical outcomes from a predefined list (e.g. symptom improvement, functionality, quality of life, remission rates, response rates, and recovery). **RESULTS:** Of 2,221 RCTs identified, 271 progressed to data extraction; 225 assessed pharmacological interventions and 46 non-pharmacological interventions. Approximately 76 outcomes were measured across the trials. The most common scale used in pharmacological trials was the Positive and Negative Syndrome Scale (PANSS) total score (76.9%), and the PANSS positive subscale in non-pharmacological trials (50%). However, even within the common outcomes, the specified level of reduction to define a relevant response varied; among trials reporting PANSS total, five different levels of reduction were defined (≥20%, ≥25%, ≥30%, ≥40%, ≥50%). Common outcomes were also measured differently in terms of improvement from baseline and proportion achieving response/remission, with little consensus on clinical meaningfulness. **CONCLUSIONS:** The RCTs included in this review reported a broad range of outcomes, making comparison of different therapies a complex task. The disparity in outcomes between pharmacological and non-pharmacological outcomes scales highlights the challenges in designing trials to demonstrate clinical benefit.

## PRM5
## MULTI-DIMENSIONAL CAPTURE OF PATIENT-RELEVANT ENDPOINTS IN REGULATORY TRIALS AND HEALTH TECHNOLOGY ASSESSMENTS IN ONCOLOGY TWO YEARS AFTER INTRODUCTION OF THE GERMAN AMNOG HEALTH CARE REFORM

Dintsios CM[1], Knoerzer D[2], Duenne AA[2], Schwartz FW[3], Ruof J[2]
[1]*German Association of Research-based Pharmaceutical Companies (vfa), Berlin, Germany,* [2]*Roche Pharma, Grenzach-Wyhlen, Germany,* [3]*Hannover Medical School, Hannover, Germany*

**OBJECTIVES:** With the introduction of AMNOG in January 2011, an early benefit assessment (EBA) was required for new medicines in Germany. EBAs are based on the additional therapeutic benefit of a drug on patient-relevant endpoints (PREs). We compared the acceptance of PREs for oncology in regulatory trials, and in EBAs conducted by German health technology assessment (HTA) bodies. **METHODS:** EBAs on oncology drugs and the respective regulatory trials were reviewed. The Federal Joint Committee (G-BA) website was used to obtain manufacturers' value dossiers, Institute for Quality and Efficiency in Health Care (IQWiG) assessments, and G-BA resolutions. Acceptance of endpoints in the dimensions of mortality, morbidity and quality of life (QoL) by HTA bodies, IQWiG and G-BA, were compared to those accepted for regulatory trials. Data on endpoints used in regulatory trials were obtained from the manufacturers' value dossiers. **RESULTS:** Overall survival (OS) and measures of disease morbidity, such as progression-free survival (PFS), were generally accepted in regulatory trials. OS was accepted by IQWiG and G-BA as a mortality endpoint for evaluating additional benefit. Widely accepted morbidity endpoints such as PFS were not deemed patient-relevant by IQWiG and G-BA. In general, QoL questionnaires used in regulatory trials were accepted by the HTA bodies, although minor variability between questionnaires led to some exclusions from the HTA evaluations and the obtained QoL data revealed a number of missing values. **CONCLUSIONS:** HTA and regulatory bodies largely agree on the acceptance of mortality and QoL endpoints typically evaluated in oncology. Considerable variability was observed in the acceptance of PREs in morbidity. Evaluating additional benefit based only on mortality and QoL endpoints underestimate the potential value of new drugs. Multiple endpoints, which capture all three dimensions, should be evaluated in regulatory trials and accepted by IQWiG and G-BA to confirm patient-relevant additional benefit.

## PRM6
## THRESHOLD SELECTION IN BIOMARKERS USING COX REGRESSION. AN APPLICATION TO NON-SMALL-CELL LUNG CANCER

Luaces P[1], Sánchez L[1], Ballesteros J[2], García B[1], Rodríguez PC[1], Viada C[1], Popa X[1], Crombet T[1], Lage A[1]
[1]*Center of Molecular Immunology, Havana, Cuba,* [2]*University of the Basque Country UPV/EHU, Leioa, Spain*

**OBJECTIVES:** To select thresholds for predictive biomarkers using Cox regression. **METHODS:** We used data from a Cuban trial designed to assess the efficacy of immunotherapy against the epidermal growth factor (EGF) to test our approach. The trial included 122 patients diagnosed with non-small-cell lung cancer (NSCLC) who had basal EGF concentration available. The EGF concentration was analysed as a predictor of immunotherapy success over the range of all possible values of the biomarker (w [a,b]). For each $w_i$, patients with $w > w_{i0}$ were selected and a Cox model adjusted to assess survival. We then identified the $w_{0s}$ with significant treatment results to find (a) the lowest biomarker threshold where the effect of treatment was significant and also to find (b) the biomarker threshold that reflected the highest difference between treatments. **RESULTS:** For NSCLC we observed that EGF concentration thresholds range from 870 pg/ml to 2000 pg/ml were significant. At the lowest threshold (870 pg/ml) the immunotherapy group showed a 6-month difference for the median survival ($p = 0.022$) whereas at the threshold that showed the maximum difference between treatments (EGF = 1750 pg/ml) the immunotherapy group presented a 10-month difference for the median survival ($p = 0.004$). **CONCLUSIONS:** The evaluation of p-values of the effect of treatment for each $w_0$ [a,b] allows the selection of the thresholds where the treatment result is significant. Whereas the

lowest threshold where the effect of treatment is significant allows the selection of patients that could be mostly benefited with the treatment, the selection of the threshold with the minimum p-value will reflect the higher difference between treatments.

PRM7

SYSTEMATIC LITERATURE REVIEW AND VALIDITY EVALUATION OF THE EXPANDED DISABILITY STATUS SCALE (EDSS) AND THE MULTIPLE SCLEROSIS FUNCTIONAL COMPOSITE (MSFC) IN PATIENTS WITH MULTIPLE SCLEROSIS

Meyer-Moock S[1], Feng YYS[1], Dippel FW[2], Kohlmann T[1]
[1]*University Medicine Greifswald, Greifswald, Germany,* [2]*Sanofi, Berlin, Germany*

**OBJECTIVES:** There are a number of instruments to describe severity and progression of multiple sclerosis, which are increasingly used as endpoints to assess the effectiveness of therapeutic interventions. We examined to what extent the psychometric properties of the two accepted instruments - EDSS and MSFC – meet the methodological standards and what value they have in clinical trials. **METHODS:** We conducted a systematic literature search in relevant databases [MEDLINE (PubMed), ISI Web of Science, EMBASE, PsycINFO & PSYNDEX, CINAHL] yielding 3,860 results. The identification of relevant full-text publications was conducted using abstract and then full-text reviews. **RESULTS:** For evaluation of psychometric properties (validity, reliability, sensitivity of change) of EDSS and MSFC, 120 relevant full-text publications were identified, 54 of them assessed the EDSS, 26 the MSFC and 40 included both instruments. The EDSS has some documented weaknesses in reliability and sensitivity to change. For the MSFC, the main limitations are the learning effects and the z-scores method used to calculate the total score. However, the methodological criterion of validity applies sufficiently for both instruments. For use in clinical studies, we found that the EDSS has been preferred as a primary and secondary outcome measure in recent studies (50 EDSS, 9 MSFC). **CONCLUSIONS:** Recognizing their strengths and weaknesses, both EDSS and MSFC are suitable to detect the effectiveness of clinical interventions and to monitor the disease progress. Almost all publications identify the EDSS as the most widely used tool to measure disease outcomes in clinical trials. Despite some limitations, both instruments are accepted to generate "hard endpoints". In no publication, EDSS or MSFC are discussed as surrogate parameters. A great advantage of the EDSS is the international acceptance (e.g. by EMA) as a primary endpoint in clinical trials and its broad use in trials, enabling cross-study comparisons.

PRM8

IMPACT OF MEDICATION ADHERENCE ON HEALTH CARE COST IN ASTHMA

Nagy B[1], Tóth E[2], Komáromi T[2], Lang Z[2], Molnár D[3], Kalmár J[3], Marusakova E[4], Helbich M[5]
[1]*Healthware Ltd., Budapest, Hungary,* [2]*Healthware Consulting Ltd., Budapest, Hungary,* [3]*GSK Hungary, Budapest, Hungary,* [4]*GlaxoSmithKline Slovakia, Bratislava, Slovak Republic,* [5]*Caldera s.r.o., Slovakia, Banská Štiavnica, Slovak Republic*

**OBJECTIVES:** To evaluate the impact of medication adherence on health care utilization and costs of the patients with asthma in Hungary. **METHODS:** The authors conducted a retrospective observation of the patients continuously enrolled in medical and prescription benefit plans from July 2007 to June 2012. The study is based on patient attendance data of Hungarian National Health Insurance Fund -NHIFA. The accessible resource uniquely contains the detailed provision data (medicine, out- and inpatient services) about the whole 10 millions Hungarian populations. Inclusion criterion for the patients was at least one diagnosis of asthma in inpatient or outpatient care (ICD code J45) and at least one relevant asthma therapy prescription in a twelve months period, and at least one relevant asthma therapy prescription during the following twelve months period. Disease-related and all-cause related medical costs, drug costs, and hospitalization risk were measured. These measures were modeled at varying levels of medication adherence using regression analysis. Adherence (Sokol, 2005) was defined as the percentage of days during the analysis period that patients had a supply of 1 or more maintenance medications for the condition. The days of supply are calculated based on WHO DDD's. This measurement strategy reduces the risk of overestimating adherence. For prescriptions extending beyond the end of the analysis period, days' supply is truncated at the end of the period. Patients in each study sample are stratified into 5 categories based on their adherence score: 1–19%, 20–39%, 40–59%, 60–79%, or 80–100 %. **RESULTS:** High level of medication adherence was associated with lower hospitalization and exacerbation rates. **CONCLUSIONS:** Increased drug utilization can provide a net economic return when it is driven by improved adherence.

PRM9

IMPACT OF MEDICATION ADHERENCE ON HEALTH CARE COST IN COPD

Nagy B[1], Tóth E[2], Komáromi T[2], Lang Z[2], Molnár D[3], Kalmár J[3], Helbich M[4], Marusakova E[5]
[1]*Healthware Ltd., Budapest, Hungary,* [2]*Healthware Consulting Ltd., Budapest, Hungary,* [3]*GSK Hungary, Budapest, Hungary,* [4]*Caldera s.r.o., Slovakia, Banská Štiavnica, Slovak Republic,* [5]*GlaxoSmithKline Slovakia, Bratislava, Slovak Republic*

**OBJECTIVES:** To evaluate the impact of medication adherence on health care utilization and costs of the patients with COPD in Hungary. **METHODS:** The authors conducted a retrospective observation of the patients continuously enrolled in medical and prescription benefit plans from July 2007 to June 2012. The study is based on patient attendance data of Hungarian National Health Insurance Fund - NHIFA. The accessible resource uniquely contains the detailed provision data (medicine, out- and inpatient services) about the whole 10 millions Hungarian populations. Inclusion criterion for the patients was at least one diagnosis of COPD in inpatient or outpatient care (ICD code J44) and at least one relevant COPD therapy prescription in a twelve months period, and at least one relevant COPD therapy prescription during the following twelve months period. Disease-related and all-cause related medical costs, drug costs, and hospitalization risk were measured. These measures were modeled at varying levels of medication adherence using regression analysis. Adherence (Sokol, 2005) was defined as the percentage of days during the analysis period that patients had a supply of 1 or more maintenance medications for the condition. The days of supply are calculated based on WHO DDD's. This measurement strategy reduces the risk of overestimating adherence. For prescriptions extending beyond the end of the analysis period, days' supply is truncated at the end of the period. Patients in each study sample are stratified into 5 categories based on their adherence score: 1–19%, 20–39%, 40–59%, 60–79%, or 80–100 %. **RESULTS:** High level of medication adherence was associated with lower hospitalization and exacerbation rates. **CONCLUSIONS:** Increased drug utilization can provide a net economic return when it is driven by improved adherence.

PRM10

USING SATURN PLOTS TO DESCRIBE CO-MORBIDITY PATTERNS WITHIN COHORTS

Juneau P
*Truven Health Analytics, Boyds, MD, USA*

**OBJECTIVES:** It is a common practice in outcomes research studies to examine several co-morbidities over two or more cohorts to develop some intuition about the health status of each group. It is also a common practice in claims data to record the co-morbid condition as a binary variable (i.e., absence or presence of the co-morbidity). If an investigator is interested in studying the absence or presence of 10 co-morbidities, he or she will have to consider $2^{10}$= 1024 possible co-morbidity patterns, a seemingly daunting task. The common current practice is to construct summary tables and examine them to understand cohort co-morbidity patterns. Even with this summarization, it is difficult to deduce what the composition is for the cohort over all co-morbidities simultaneously. Once again, this can be a daunting task. The objective of this research is to develop a means of summarizing these rich and somewhat complex data to enhance clinical decision making. **METHODS:** Graphical approaches for the summarization of data enable geometry, scaling, shading, and or color to describe such "high dimensional" data. The author will introduce a novel means of plotting the co-morbid conditions that will afford investigators the ability to study patterns of co-morbidities simultaneously and understand the relative frequencies of their occurrence in one display. **RESULTS:** The use of a novel graphical procedure (a Saturn plot) allows an investigator to examine co-morbidity patterns readily when the number of binary co-morbidities is 10 without having to resort to poring over several tables or one large table partitioned into smaller ones based on co-morbidities. **CONCLUSIONS:** A newly developed graphical data summary called a Saturn plot allows investigators to indentify the relative frequency of various subgroups (as defined by their co-morbidity pattern) within a cohort without the need to study large sets of tables.

PRM11

DESIGN OF A RANDOMIZED CONTROLLED TRIAL (RCT) EVALUATING OUTCOME AND COST-EFFECTIVENESS OF A LOCAL CASE MANAGEMENT INTERVENTION OF PATIENTS SUFFERING FROM CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

Sørensen SS[1], Pedersen KM[2], Ehlers L[1]
[1]*Aalborg University, Aalborg, Denmark,* [2]*University of Southern Denmark, Odense M, Denmark*

**OBJECTIVES:** In December 2011 the Danish Government issued a new plan of action for chronic disease management in the Danish counties and DKK 100 million were granted to set up new positions as case managers to help vulnerable elderly patients. No precise job description was provided, however, and the Danish counties request evidence for the effect of case management (CM). The aim of this study was to 1) design a job description for a case manager, and 2) design a RCT evaluating consequences and costs of providing local CM to patients with COPD. **METHODS:** By use of the UK Medical Research Councils (MRC) framework for development of complex interventions, the design of the case manager job description and the RCT was determined through a systematic literature review, interviews with key persons and discussions in a specialist-comprised steering group. **RESULTS:** CM was designed to encompass coordination of care, facilitation of relevant health- and social services and promotion of patient self-care through advocacy and education. The RCT was powered to detect the effect of CM on hospital admissions. Secondary measures include mortality, quality of life, self-care and cost-effectiveness of CM versus usual care. 150 COPD patients are randomized into two groups after referral to pulmonary rehabilitation at the local rehabilitation center in Aalborg County, Denmark. The control group will receive usual care, whereas the interventional group will receive CM besides their usual care. Each patient is followed for 12 months. The questionnaires SF-12, EQ-5D, Sct. George-Respiratory-Questionnaire (SG-RQ) and The Patient-Activation-Measure (PAM-13) are completed at baseline and 12 months. Prospectively collected data from national population-based medical registries are used to estimate events and resource usage. **CONCLUSIONS:** The study is expected to provide further insight to the future organization of CM, and if being cost-effective, the intervention could be applied to comparable health care settings.

PRM12

REVIEW OF META-ANALYSIS METHODS FOR MULTINOMIAL DATA

Bennison C[1], van Beurden-Tan C[2], Buyukkaramikli C[2], Heeg B[2]
[1]*Pharmerit Ltd., York, UK,* [2]*Pharmerit International, Rotterdam, The Netherlands*

**OBJECTIVES:** Indirect comparisons are often based on binary outcomes (e.g. relapse / remission) or continuous outcomes. In these cases logistic or linear models are applied to make the indirect comparisons. However, sometimes datasets contain multinomial outcomes, such as 'complete', 'partial' and 'no' response in oncology, that need to be indirectly compared. With multinomial data, different indirect comparisons methods may be required to answer different research questions. Our goal was to identify and qualitatively compare the different techniques that have been used to model multinomial data in an indirect comparison framework. **METHODS:** A systematic review of the PubMed database was conducted to identify different methods for handling multinomial data in a meta-analysis. Key words included 'meta-analysis', 'ordinal', 'ordered', 'multinomial' and 'proportional odds', in various combinations. Models were qualitatively compared according to their assumptions, flexibility and complexity. **RESULTS:** The systematic review identified three methods: a proportional odds model, an ordered logistic model, and a multinomial model. The proportional